

Университет ИТМО

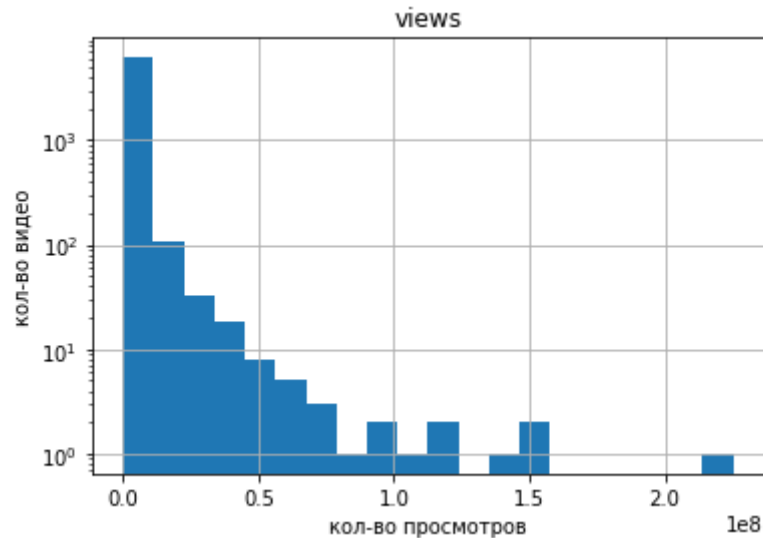
Практическая работа №2
по дисциплине «Визуализация и моделирование»

Автор: Сулейманов Руслан Имранович
Поток: 1.1 **Группа:** к3241 **Факультет:** ИКТ
Преподаватель: Чернышева А.В.

Санкт-Петербург, 2021 г.

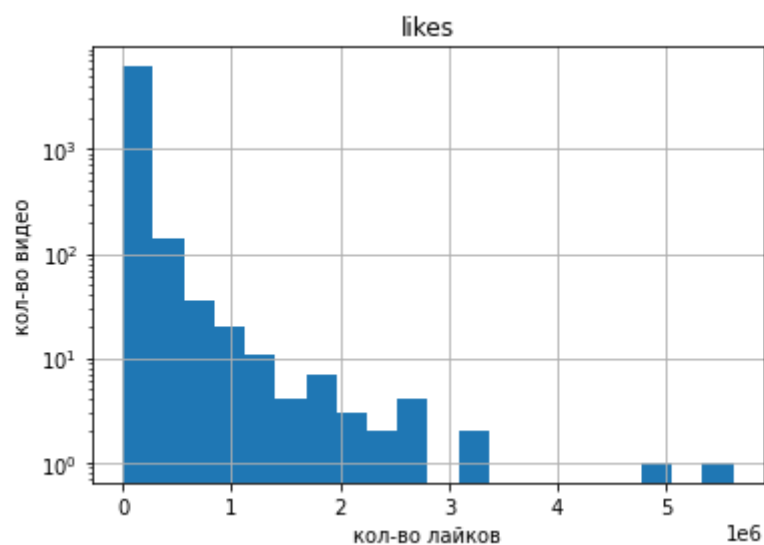
Визуализация данных

1. Посмотрим на соотношение количества видео и количества их просмотров для собирательной статистики.



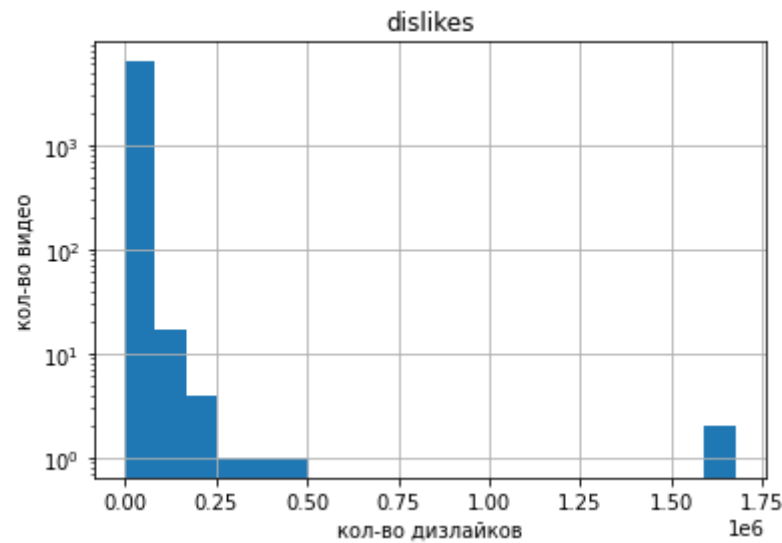
Из графика мы видим, что количество просмотров убывает по экспоненте, что логично, мало видео имеют много просмотров. Но, весьма странно, что мы имеем пропуск среди значений $1,7 \cdot 2 \cdot 10^8$ просмотров. Вероятно, слишком мало видеороликов могут достигнуть большого количества просмотров, а те что достигли, набирают еще больше популярности

2. Посмотрим на соотношение количества лайков и количества их просмотров для собирательной статистики.



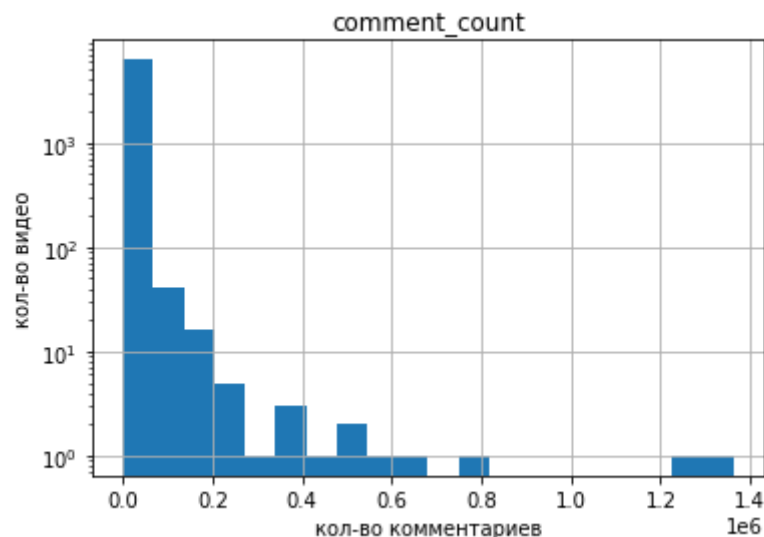
Аналогично, убывает по экспоненте, но при этом все также имеется пропуск значений.

3. Посмотрим на соотношение количества дизлайков и количества их просмотров для собирательной статистики.



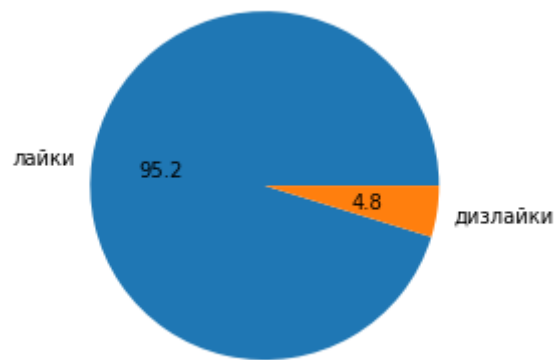
Аналогично, убывает по экспоненте. Также множество видео не набирают много дизлайков, что говорит нам о том, что люди редко ставят дизлайки, и предпочитают просто игнорировать видео-ролики, но при этом, у некоторых видеороликов огромное число дизлайков, что означает, что некоторые видео просто ужасны, и люди не могут пройти мимо.

4. Посмотрим на соотношение количества комментариев и количества их просмотров для собирательной статистики.



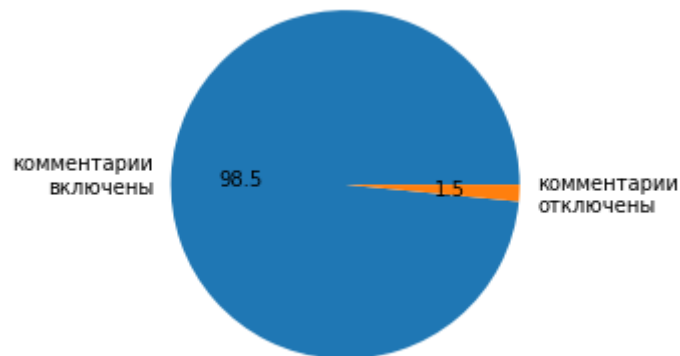
Тут все весьма аналогично лайкам, что вполне логично.

5. Посмотрим на соотношение лайков и дизлайков на видеороликах в процентах.



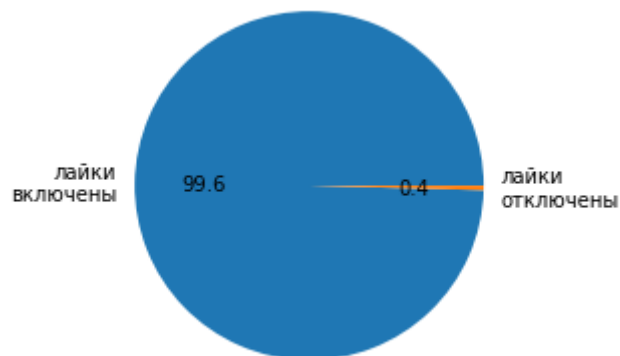
Еще раз убедились, что люди предпочитают поставить лайк, нежели дизлайк

6. На каком процентов видеороликах отключены комментарии



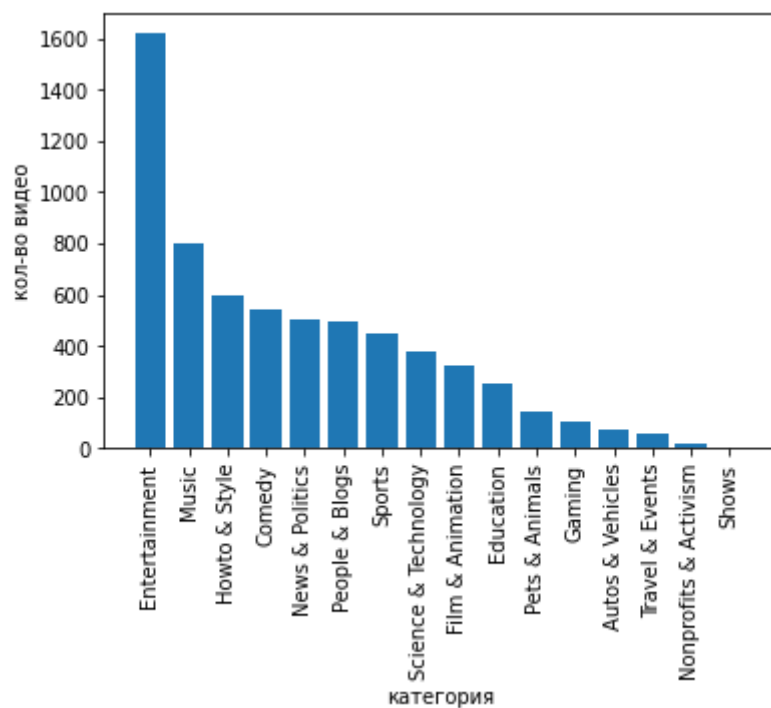
Только 1,5 процентов, подавляющее число людей предпочитают получать фидбек от пользователей. А что же с лайками?

7. На каком процентов видеороликах отключены лайки



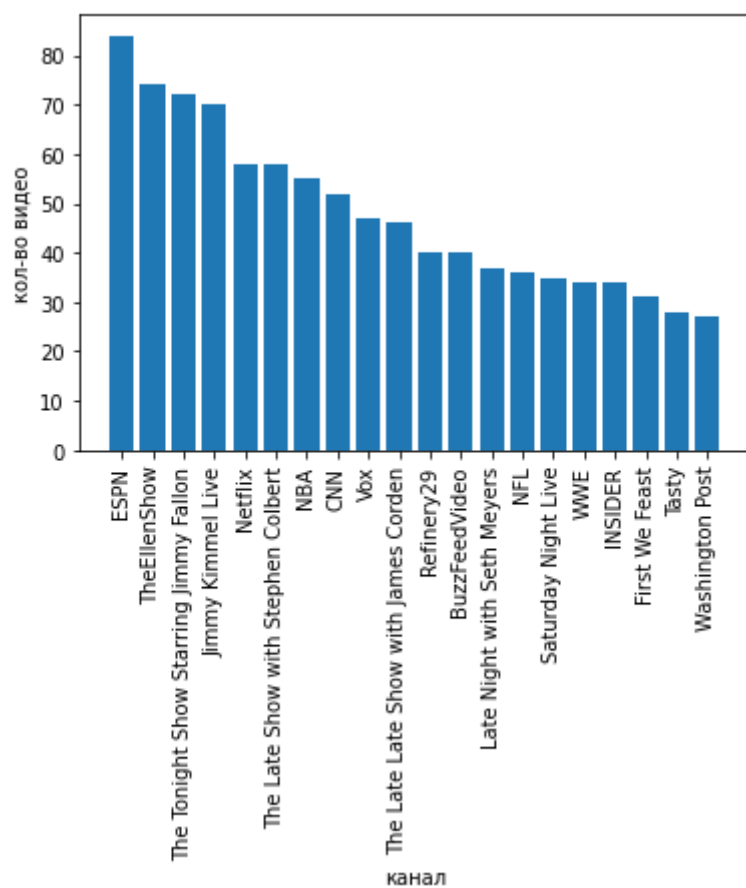
Еще меньше, что и логично, зачем отключать положительный отзыв о видеоролике?

8. Теперь давайте посмотрим на топ категорий



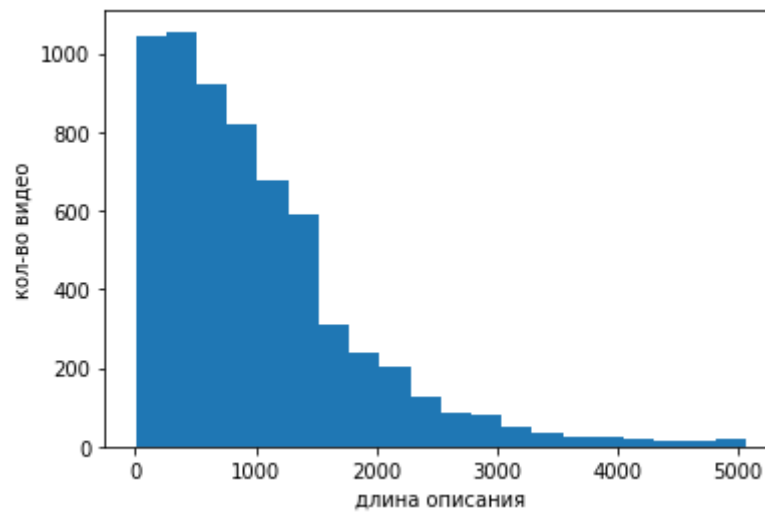
Развлечения набирают самую большую популярность. Вероятно, потому что в последнее время очень распространены тиктоки, которые не занимают много времени, и дают людям расслабиться, вскружиться или улыбнуться.

9. топ каналов, чаще всего попадавших в тренды



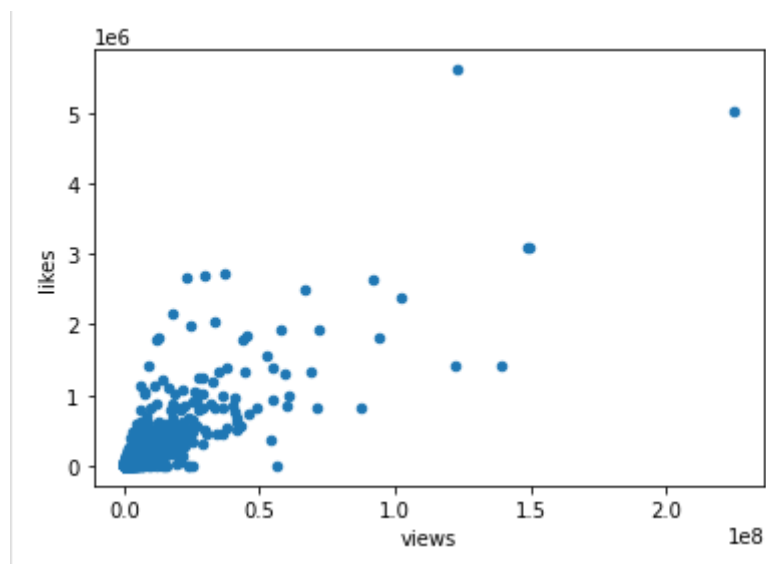
Неудивительно, что на первом месте развлекательные каналы.

10. Распределение длин описания (в символах)



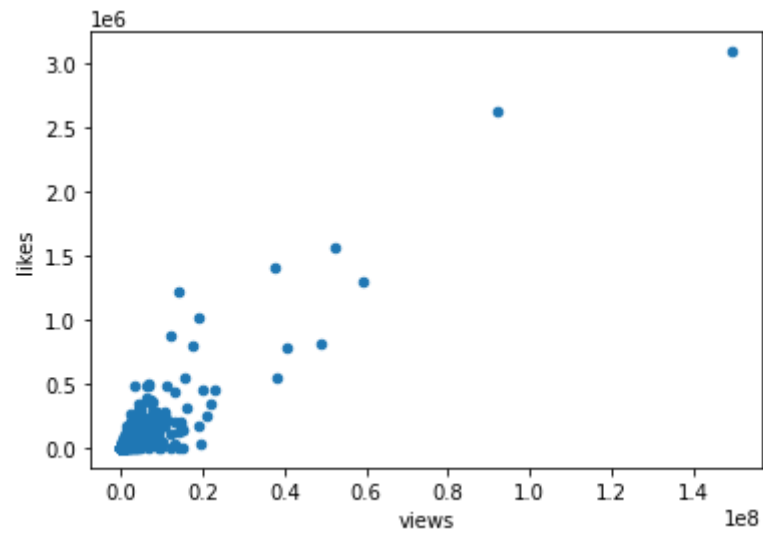
Из графика видна гиперболическая зависимость, но с резким спадом. Это можно объяснить клешированными комментариями, не занимающих много слов. Редко люди оставляют большой отзыв о видеоролике.

11. Теперь сравним отношение количества лайков к количеству просмотров.



Из графика ничего не понятно, попробуем сузить поиск.

12. Отношение количества лайков к количеству просмотров для видео категории Entertainment (id = 24)



Понятнее не стало, видна более менее параболическая зависимость, но все еще не понятно что творится в самом начале, так как количество видеороликов слишком большое.

Выводы: построены разные зависимости для анализа и описательной статистики датасета. Проанализированы графики. Код предоставлен в файле lab2.ipynb